

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G11B		A2	(11) International Publication Number: WO 99/03098 (43) International Publication Date: 21 January 1999 (21.01.99)
(21) International Application Number: PCT/US98/14247 (22) International Filing Date: 9 July 1998 (09.07.98) (30) Priority Data: 60/052,055 9 July 1997 (09.07.97) US 09/111,691 8 July 1998 (08.07.98) US (71) Applicant: SUTMYN STORAGE CORPORATION [US/US]; P.O. Box 5687, Santa Clara, CA 95056-5687 (US). (72) Inventors: YATES, Neville; 19644 Montevia Road, Los Gatos, CA 95030 (US). MILLER, Jeffrey; 133 Spruce Avenue, Menlo Park, CA 94025 (US). BOUSSINA, Touraj; 316 Clifton Avenue, San Carlos, CA 94070 (US). HARANO, Allen; 5672 Poglia Court, San Jose, CA 95136 (US). (74) Agents: BECK, David, G. et al.; Townsend and Townsend and Crew LLP, 8th floor, Two Embarcadero Center, San Francisco, CA 94111-3834 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: IMPROVED INTERFACES FOR AN OPEN SYSTEMS SERVER PROVIDING TAPE DRIVE EMULATION (57) Abstract An improved interface between a host computer and a tape drive e mulation system includes software interfaces for communicating control, configuration, and policy data and a hardware interface for providing redundancy and fan-out between the main controller and host channels.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon	KR	Republic of Korea	PL	Poland		
CN	China	KZ	Kazakhstan	PT	Portugal		
CU	Cuba	LC	Saint Lucia	RO	Romania		
CZ	Czech Republic	LI	Liechtenstein	RU	Russian Federation		
DE	Germany	LK	Sri Lanka	SD	Sudan		
DK	Denmark	LR	Liberia	SE	Sweden		
EE	Estonia			SG	Singapore		

IMPROVED INTERFACES FOR AN
OPEN SYSTEMS SERVER PROVIDING TAPE DRIVE EMULATION
5 CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority from U.S. Provisional Application No. 60/052,055, filed July 9, 1997, which is incorporated herein by reference in its entirety for all purposes.

10

BACKGROUND OF THE INVENTION

The present invention relates generally to data storage systems and more particularly relates to tape drive emulation (TDE) systems.

15

Many data processing systems utilize tape drives for storage of data. Channel interfaces and commands generated by a data source to control the transfer of data from a host computer to a tape drive are well-known in the art and will not be described in detail here. One example of an operating system for managing data transfer between a host computer and a tape drive is the MVS system manufacture by IBM Corporation.

20

In a tape drive emulation system, the data output by the host is not actually written to a tape drive and data input by host is not actually read from a tape drive. Instead, in one type of TDE system, the data is input from and output to staging disks.

25

SUMMARY OF THE INVENTION

30

According to one aspect of the invention, an improved interface facilitates communication between a library management system (LMS), operating on a host computer, and an Open Systems Server (OSS) containing virtual tape drives (VTDs) and operating as a TDE system.

35

According to another aspect of the invention, the interface supports two different kinds of communications. One type is a dump of a large amount of data that is not suitable to real-time short transactions. In this case, the data to be provided to the LMS is packaged in a virtual volume with a

special name convention. Such a special virtual volume is called an "administrative volume." Periodically, the LMS requests an administrative volume, the OSS "mounts" the volume on a VTD, and the LMS reads the wanted information from the 5 VTD. Thus, the administrative volume is used to communicate status, control, and configuration information between the LMS in the host and OSS, using a standard access method (tape or virtual tape).

According to another aspect of the invention, the 10 interface utilizes load display (LD) commands. LD commands are channel commands normally used to route messages to a tape drive's operator display accessory. The present invention utilities LD commands to communicate policy and control 15 information messages to the TDE system and to monitor its condition. Policy information reflects decisions by the user(s) and includes rules and preferences for the handling and disposition of data written to VTDs. These user decisions are unknown to OSS, but are communicated to the LMS in the host, and guide LMS in its operation. For example, the 20 assignment by LMS of particular data to a virtual volume belonging to a virtual volume set (VSET) with which a collection of pre-programmed handling rules is associated is an expression of policy. The virtual volumes need a large 25 amount of policy info for management.

According to another aspect of the invention, the SBUS card slots are expanded to facilitate the use of a SPARC CPU to control a mass storage system.

According to another aspect, a special "Health Check" LD message is periodically sent. If a critical 30 situation exists within OSS, a special error message will be generated and delivered to the operator by LMS.

According to another aspect of the invention, SBUSS of two SPARC CPUs are connected to each ESCON interface for redundancy.

Additional features and advantages of the invention 35 will be apparent in view of the following detailed description and appended drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1A is a block diagram of a preferred embodiment of the invention;

5 Fig. 1B is a block diagram of a data storage system including a TDE system;

Fig. 2 is a flow-chart describing the steps of mounting an administrative volume;

Fig. 3 is a perspective view of the channel interface hardware;

10 Fig. 4 is a block diagram of a DSB board and an ESCON Interface daughter card;

Fig. 5 is a perspective view of the channel interface hardware coupled to primary and alternate main processors; and

15 Fig. 6 is a block diagram depicting redundant connections to the CIFs.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

A preferred embodiment will now be described with reference to the figures, where like or similar elements are designated with the same reference numerals throughout the several views. Fig. 1A is a high level block diagram of a part of a tape drive emulation (TDE) system 10, also referred to herein as OSS 10, utilizing an embodiment of the present invention.

A plurality of channel interfaces (CIFs) 12 are coupled to host I/O channels (not shown) to transfer data between the host and the TDE system.

Each CIF 12 includes a host interface 14, an embedded controller 16, a data formatter 18 for performing data compression and other functions, an SBUS interface 22, a buffer memory 20, and an internal bus 24. In the preferred embodiment, the embedded processor 16 is a model i960 manufactured by Intel Corporation.

35 The main controller 30 includes a main processor 32, main memory 34, an SBUS interface 36, and an internal bus 38. In the preferred embodiment, the main processor is a SPARC computer manufactured by Sun Microsystems Incorporated. The

CIFs 12 and main controller 30 are coupled by a system bus (Sbus) 40.

The tape drive emulation (TDE) system 10 stores host data on "virtual tape drives." In one preferred embodiment, 5 the data is actually stored on staging disks. Because the TDE system 10 must interact with the host as if the data were actually stored on tape drives, a data structure called a virtual tape drive (VTD) is maintained in main memory 34 for each virtual tape drive. Each VTD contains all information 10 about the state of the associated virtual tape drive.

Fig. 1B is a high-level block diagram of a system in which a preferred embodiment of the invention is utilized. In Fig. 1B, a host computer 50, for example an IBM mainframe computer, executes a plurality of applications 52.

15 In practice, the host computer 50 is typically running the MVS operating system manufactured by IBM. MVS provides the applications with I/O services, including I/O to an automatic tape library (ATL) 54. The physical interface between the applications 52 and ESCON tape drives 55 is the 20 ESCON 3490 magnetic tape subsystem interface 55a. MVS, the ESCON interface 55a, and the host computer 50 are well-known and not a part of the present invention.

The preferred embodiment of tape drive emulation (TDE) system 10, designated OSS 10 (open systems server), is 25 manufactured by the assignee of the present invention. OSS 10 maintains virtual tape drives 56 (VTDs) which emulate the physical ETDs 55. More details of the VTDs 56 will be presented below. The interface between an application 52 and a VTD 56 is the OSS Emulated Device interface 57.

30 A library management system (LMS) software module 60 resides on the host 50 and provides services to MVS and OSS. LMS 60 is responsible for management of the tape library environment and performs such tasks as fetching and loading cartridges into drives, returning unloaded cartridges to their 35 home locations, etc. The interface between LMS 60 and OSS 10 is the library manager interface, with paths 62a and 62b based on two distinct protocols.

The VTD 56 is a non-physical device that responds as if it were a real device. In the currently described embodiment, the emulated physical device is an IBM-3490 tape drive. The VTD 56 responds to commands issued on a channel in the same fashion as the emulated technology.

Host data is stored in volumes. A virtual volume is a collection of data and metadata that, taken together, emulate a real tape volume. When "mounted" on a VTD, these virtual volumes are indistinguishable from real tape volumes by the host computer.

In this context "data" refers to data output by the host to be stored on tape and "metadata" refers to information generated by OSS which permits the emulation of real tape drives and volumes.

An example will help clarify the meaning of the terms. If a host application intends to write data to tape, it requests that a tape be mounted on a tape drive. LMS intercepts the request and causes a virtual volume to be mounted on a virtual tape drive to receive the application output, which is delivered by the ordinary tape output programs of the operating system. Blocks of data received by OSS are "packetized," the packets are grouped together in clusters with a fixed maximum size, called "extents," and the extents are written to staging disks. Often the extents containing data from one virtual tape are scattered over several disk drives. All information about the packetization, such as packet grouping in extents and extent storage locations, required to reassemble the volume for later use by the host is metadata. Part of the metadata is stored with each extent and part is stored on non-volatile in OSS, separate from the extent storage.

LMS requires information concerning the contents of OSS to properly respond to host requests for accessing VTDs and virtual volumes. It also needs information on OSS storage space usage to manage auxiliary operations which maintain enough free space to adequately receive new outputs. In the present embodiment, there are two primary interfaces for transferring information between LMS and OSS. In the

discussion below, an interface means a protocol or style of interaction between LMS and OSS, not necessarily its physical implementation.

The first interface to be described is the
5 administrative volume interface, which is used to access a large volume of information during a single transaction. This type of information relates to complete and detailed status of the virtual library in OSS and includes more information than can be rapidly transferred using short messages. The high
10 level description of the function of the administrative volume interface will now be presented with reference to Fig. 2.

When LMS requires information describing the status and contents of OSS, it uses the conventional facilities of the operating system to allocate a VTD and request the
15 mounting of an administrative volume. For example, it is necessary to periodically synchronize status and content information of OSS with information in the tape management component in use by the operating system. Utilities in LMS implement this synchronization.

20 A special naming convention for the administrative volumes allows OSS to interpret the mount command as a request for a particular body of status information. Different names specify different types of administrative volumes.

These administrative volumes appear to host
25 applications as ordinary volumes with IBM standard labels. According to a convention used in the preferred embodiment, their volume serial numbers, the names by which they are known in host indexes of tapes and which are written on the tapes in the VOL1 labels, are generated by adding a single character to
30 a specified five character prefix, reserved to the exclusive use of LMS.

In the currently described embodiment, there are five types of administrative volumes defined: Audit List; Audit List Discrepancies; Audit List Agreement; RAID Status
35 Report; and OSS Data Dump. Presently, only read-only administrative volumes, i.e., administrative volumes transferring information from OSS to LMS, are implemented. Each volume contains labels and one data set.

The "Audit List" data set contains a volume status record for each volume known to the OSS. This data set is a read-only data set produced by the OSS when the appropriate administrative volume is mounted.

5 The "Audit List Discrepancies" data set also contains volume status records. It is a write-only data set and will contain a volume status record for every volume on which a host's tape management system component and OSS differ on volume status.

10 The "Audit List Agreement" data set also contains volume status records. It is a write-only data set, and will contain a volume status record for every volume on which the tape management system component and the OSS agree on volume status.

15 The "RAID Status Report" data set contains a virtual volume set (VSET) usage record for each VSET known to OSS, a free space record for each RAID region and a single OSS system status record. This data set is a read-only data set produced by the OSS when the appropriate administrative volume is
20 mounted.

25 Information in the "OSS Data Dump" data set will contain raw data that the OSS wishes to communicate to a host application. This data set is a read-only data set produced by the OSS when the appropriate administrative volume is
mounted.

30 In response to the host's mount request for an administrative volume containing a read-only data set, OSS builds an administrative data set from status information stored in its data base. The type of information included in
the administrative data set depends on the name (the volume serial number) included in the mount command.

35 When OSS has completed building the administrative data set and storing it on disks in the usual way for storage of OSS virtual volumes, it signals LMS that the administrative volume is mounted. LMS then reads the administrative data set to obtain the requested status information.

Thus, standard channel commands are utilized to transfer status information between OSS and LMS in an efficient manner.

5 The above-described administrative volume interface provides a unique mechanism to transfer large volumes of status, control, and configuration information between OSS and LMS. Such information is required in a TDE system to permit effective management and operation of the TDE system by the LMS.

10 Another type of information unique to a TDE system is termed "policy" information. In one type of OSS, virtual volume data transferred between the host is staged on disk drives. Additionally, tape drives (the SCSI tape drives of Fig. 1) are also included and virtual volumes may be destaged 15 from the disk drives to the tape drives.

20 The majority of optimizations available to tailor an OSS to particular customer's requirements come from optimization of timing of various events in the life-cycle of a virtual volume. These optimizations take the form of choices among various policies defined below.

The task of managing policy decisions is simplified by grouping attributes of virtual volumes as follows:

- those attributes which specify the kind of medium being emulated;
- 25 -- those attributes which guide the choice of long term storage media for data associated with the virtual volume; and
- those attributes that direct the timing of data residency as it passes through the OSS staging disks.

30 Several examples of policies associated with a virtual volume are the performance class, media class, and storage class.

The performance class specifies the attributes of a 35 virtual volume that govern residency timing of the data of the virtual volume on various media. It is called performance class because altering the class changes user-perceived performance, most notably mount delay.

The media class describes the attributes of a single kind of media emulated by OSS, i.e., attributes such as technology type, media geometry, etc. Media classes are defined by the user and associated with a virtual volume at 5 the time of its creation. An example of media class might be "3490-CST," defined as 550 foot, 3490 EDRC tapes.

The storage class is a description of whether data is replicated, and how data is stored in OSS. A storage class is associated with a virtual volume at the time of its 10 creation, and may be changed at any time that a volume is mounted as an output-only volume. Storage classes are defined by the user. An example of a storage class might be "VAULTED" defined to direct the data for a virtual data to a single stacked image and single native image, with the intent that 15 the native image will be stored offsite.

As these examples show, policy information must be communicated to OSS in real time, for example when a virtual volume is created or mounted. In the present embodiment, 20 policy information is communicated by unconventional use of the standard load display interface (LDI).

"Load Display" is a command issued to a 3490 tape drive which transmits an associated message, part of which is to be displayed on a display pod associated with the tape drive. In an IBM 3490, the messages do not affect the 25 operation of the tape drive, but were intended to communicate with a human operator. With the advent of automatic tape libraries the LDI has become a vestigial communication channel. In the present invention, the LDI is used to communicate policy information to the OSS from the LMS in real 30 time.

A load display command transfers sixteen bytes of information, viewed as a message. The format utilized in the currently described embodiment includes a first byte set to identify the message as an LMS to OSS communication, a second 35 to identify a particular request, thirteen bytes of request specific data, a check sum byte (XOR of all preceding bytes), and a byte specifying the data length. The LDI messages are not intended to be displayed.

In the currently described embodiment, the LMS library driver uses the LDI to request these services:
mountVirtualScratch(); MountVirtualVolume();
keepVirtualVolume(); updateVolumeStatus(); healthCheck();
5 timeStamp(); stageVirtualVolume; destageVirtualVolume(); and
reuseSpace().

The mountVirtualScratch() request, for example, specifies a VSET Name in thirteen bytes of the message. Responding to the request, OSS mounts a volume it chooses from 10 a pool of volumes whose names are within a name range allocated to the VSET and whose present status is "scratch," meaning not in use, containing only a label. The volume so mounted takes on the media, performance and storage class attributes associated with the VSET as defaults. Subsequent 15 LDI requests naming the chosen volume may be used to alter certain of the attributes.

The LDI is, by its nature, a one way communication channel: it was designed for the host computer to send display messages to operators. However, one type of LDI message 20 supported by the present embodiment, the health check, is an example of the use of the LDI for information gathering rather than expressing policy or requesting action. The healthCheck() message format is sent from LMS as a poll request to determine the operational status of the OSS.

25 In critical situations the OSS must inform LMS and/or the operator that an event has occurred. These events are currently associated with RAID space shortage and equipment failures.

The LDI healthCheck() poll message is issued to the 30 OSS on a regular basis, e.g., every 30 seconds. The healthCheck() message contains a host system identifier and the date and time. If an error condition exists in OSS, the OSS ends the execution of the load display command with 35 command-failure status (unit check). The host reacts to unit check by reading sense bytes from the OSS and OSS includes in these a distinctive "SUU ERP code" (X'42'). The "EC level" byte in the sense data contains the OSS system error code, an unsigned, 8-bit number indicating the highest priority error

existing at the time. The OSS continues to respond to subsequent healthCheck() polls until the malfunction or other emergency is resolved (or goes away).

5 The MVS operating system produces a LOGREC record and an IOS000I message for each SUU response but retries the failing I/O. OSS, recognizing the retry, allows the command to complete without error. LMS intercepts the MVS IOS000I message, deletes it, informs the operator of the error condition, and then takes appropriate action.

10 An additional requirement of OSS is to provide multiple channel interfaces to a host or multiple hosts. The controller utilized in the currently preferred embodiment is based on a SPARC processor manufactured by SUN Microsystems.

15 A special interface has been designed to expand the standard interface provided with the SPARC computer. Additionally, redundancy is built into the interface to assure the reliability required of a tape storage system.

20 Fig. 3 is a high-level schematic diagram showing the redundant, channel interface expansion chassis hardware 300. The CIF interface cards 12 reside in the expansion chassis 300. The chassis is divided into two halves 302 and 304. Electrically, each half is separately powered and contains its own separate sets of SBUS connections. Each half can contain up to four dual SBUS base (DSB) boards 306, each of which can 25 contain up to four channel interfaces 12 (in this embodiment ESCON interface (EI) or block multiplexer interface (BMUX) daughter cards) for a total of sixteen interfaces per half. Unless otherwise specified, the term "chassis" in the following refers to a chassis half.

30 Each connection to the main processor(s) 320 is made via an SBUS expander, consisting of an SBUS adapter (SSA) 340 that plugs into an SBUS slot inside the SPARC, an interconnecting cable 342, an SBUS expander (SSE) 360, which connects to either two or four slots in the chassis. Each 35 slot in the chassis can connect to two SSEs 360 via identical connectors. If each SSE is connected to four slots, then two separate SBUS connections to SPARC(s) are provided for that chassis. If each SSE is connected to two slots, then each

pair of slots contains connections to two SBuses, for a total of four separate SBuses per chassis [half], and a system total of eight SBus connections to SPARCs.

Fig. 4 is a simplified block diagram of the DSB and CIF daughter cards shown in Fig. 3. A DSB 306 plugs into a slot in the chassis, and therefore connects to two SSEs 360, which are referred to as SBus0 and SBus 1. There are two SI64 bridge chips on the DSB, one for each SBus.

The SI64s 370 are connected to the [up to] four interface daughter cards on the DSB via a common i960-compatible bus 372. Also included on the DSB are some shared control/status/interrupt support registers used for communication between the SPARC and the i960s.

Thus, each DSB card 306 contains two SI-64's, one for connection to each of two SPARC SBuses. On the i960 side, the two SI-64's connect to a common bus. Each EI (Escon Interface which contains an i960) board 12 carried by the DSB 306 has access to this common bus according to a scheme of arbitration supported by a small amount of DSB logic. This connection uses a set of transceivers to link temporarily an i960's local bus and the common bus 372, which have the same design. Thus, each i960 can access and use either SI-64.

The SI-64 is a Motorola product which operates as a "gateway" between the main bus of an i960 (call this the "local" bus) and the I/O bus (specifically SBus) of a SPARC machine. It can be programmed to perform direct memory access (DMA) transfers between memories of the two computers. It is therefore able to arbitrate accesses to each kind of bus and has address and length counters to effect one transfer at a time. Without using the DMA feature, the chip allows an i960 program to read from or write to an arbitrary SPARC memory location and allows a SPARC program to read from or write to an arbitrary i960 memory location.

The SPARC system utilizes virtual addresses, which are translated into physical addresses on the SBus, for devices on the SBUS.

Access to the interface cards from the SBus (SPARC) is accomplished by dividing twenty-eight bit physical

addresses allotted to the chassis ("SBus Slot") into four equal 64 MByte areas that select one of the four DSB boards in the chassis. The 64 MByte area of each DSB board is sub-divided into four areas of 16 MB which select one of the 5 interface daughter cards 12. The 16 MB area assigned to each interface card is further sub-divided into an 8 MB area that allows direct access into the interface card. The remaining 8 MB area is mapped into a common area that allows access to shared resources on the DSB such as dual-ported RAM, bridge 10 registers, interrupt status, etc.

As depicted in Fig. 5, redundancy is provided by connecting a backup main processor 320(2) to a third SSE 360(3), coupled to the same DSBs 380 as the first SSE 360(1). Thus, it is possible for the backup main processor 320(2) to 15 take over the functions of the primary main processor 320(1) in the event of a failure.

The redundant connection the Sbuses of the primary and backup main processors is depicted in Fig. 6. Each DSB contains up to four CIFs 12 (in this figure designated ESCON 20 Interface Daughter Cards (EIs)).

The invention has now been described with reference to the preferred embodiments. Alternatives and substitutions will now be apparent to persons of skill in the art. For example, particular products such a SPARC processor and SI64 25 interface chips have been described. Other products may be substituted. Accordingly, it is not intended to limit the invention, except as provided by the appended claims.

WHAT IS CLAIMED IS:

1 1. A method of interfacing a library management
2 system software module to a virtual tape drive server module
3 with the library management system software module resident on
4 a host computer running an operating system (OS) and
5 applications, where the OS provides services for inputting and
6 outputting data sets to physical tape drives, with the
7 services including channel commands for outputting tape
8 mounting requests and for reading a tape, with the virtual
9 tape drive server for receiving commands from the host to
10 store and retrieve data from non-physical virtual tape
11 volumes, and with the virtual tape drive server module for
12 storing metadata describing the actual storage of the virtual
13 tape volume contents, and with the library management system
14 software module generating requests to cause the OS to issue
15 messages and commands to the virtual tape drive server module
16 to make the virtual tape drive server module status
17 information available to the library management system
18 software module, said method comprising:

19 establishing a naming convention to identify
20 differently named administrative virtual volumes for providing
21 different types of status information;

22 determining that the library management system
23 software module requires updating of a designated type of
24 status information;

25 generating, with the library management system
26 software module, a mount request including a designated name
27 of the administrative volume that provides the designated type
28 of configuration metadata;

29 responding, at the virtual tape drive server module,
30 to said mount request to retrieve said designated type of
31 status information and build an administrative data set
32 including the designated type of status information, storing
33 said data set in a virtual tape volume, and mounting the
34 volume on a specified virtual tape drive;

35 generating, with the library management system
36 software module and the OS, read commands to the said virtual
37 tape drive;

38 responding, at the virtual tape drive server module,
39 to said read commands to transfer said administrative data set
40 to the host for use by the library system software module.

1 2. A method of communicating, in real-time, policy
2 and control messages, indicating attributes such as the
3 performance class, media class, or storage class of a virtual
4 volume, between a library system software module and a virtual
5 tape drive server module, with the library system software
6 module resident on a host computer running an operating system
7 (OS) and applications, where the OS provides services for
8 inputting and outputting data sets to physical tape drives,
9 with the services including channel commands for transmitting
10 display messages to tape drives, with the virtual tape drive
11 server for receiving said messages and interpreting them as
12 requests from the library management system software module to
13 mount non-physical virtual tapes, create non-physical virtual
14 tape volumes, and other control purposes, and with the virtual
15 tape drive server module requiring policy and control
16 information when creating, mounting or disposing of a virtual
17 tape volume relating to user choices concerning virtual
18 storage, and with requests to cause the OS to issue commands
19 transmitting messages to the virtual tape drive server module
20 conveying the needed policy and control information, and with
21 the host computer including a load display interface for
22 transmitting messages to be displayed on a display pod of a
23 tape drive, said method comprising:

24 generating, using an application, a request to the
25 OS to mount a tape volume in which to store data;

26 intercepting, with the library management system
27 software module, said mount request;

28 building, with the library management system
29 software module, a load display interface message in load
30 display interface format for transmitting to the virtual tape
31 drive server module over the load display interface channel,

32 with said message including an identification parameter
33 identifying eligible virtual volumes and including parameters
34 specifying user-selected policy information relating to the
35 selected virtual volume to be mounted;

36 transmitting said load display interface message to
37 said virtual tape drive server module utilizing the load
38 display interface channel;

39 responding, at said virtual tape drive server
40 module, to said load display interface message to select a
41 virtual volume from among the specified eligible volumes and
42 mount said selected virtual volume and set policy parameters
43 of said selected virtual volume equal to the policy parameters
44 identified in said load display interface message.

1 3. The method of claim 2 further comprising:
2 determining at said virtual tape drive server module
3 that an error condition exists;
4 including a field in said load display interface
5 message identifying said message as a health check;
6 if an error condition, reporting an error condition
7 to the OS when the health check message is received.

1 4. A tape drive emulation interface, connecting a
2 plurality of host channels to a main processor having a
3 systembus of a tape drive emulation system, said interface
4 comprising:

5 a plurality of dual systembus base boards;
6 a plurality of channel interfaces disposed on each
7 dual systembus base board, each channel interface for
8 interfacing a host channel to input/output buffers of the main
9 processor;

10 a first bus expander, multiplexing said plurality of
11 dual systembus base boards to a bus expander port;

12 a first set of signal lines coupling said bus
13 expander port to said main processor systembus.

1 5. The interface of claim 4 further comprising:
2 a backup main processor having a systembus;
3 a second bus expander, multiplexing said plurality
4 of dual systembus base boards to a second bus expander port;
5 a second set of signal lines coupling said second
6 bus expander port to the systembus of said backup main
7 processor.

1 6. A method for transferring status information,
2 relating to virtual volumes in a tape drive emulation system,
3 from the virtual tape drive server to a host computer, said
4 method comprising the steps of:

5 transmitting a request for status information from
6 the host computer to the virtual tape drive server;
7 in response to said request, building and mounting
8 an administrative virtual volume at said virtual tape drive
9 server to store requested status information;
10 reading said administrative virtual volume to
11 transfer said requested status information from said
12 administrative virtual volume to said host computer.

1 7. A method for updating status information,
2 relating to virtual volumes in a tape drive emulation system,
3 from a host computer to the virtual tape drive server, said
4 method comprising the steps of:

5 transmitting a request to update status information
6 from the host computer to the virtual tape drive server;
7 in response to said request, mounting an
8 administrative virtual volume at said virtual tape drive
9 server to store updated status information;
10 writing to said administrative virtual volume to
11 transfer said updated status information to said
12 administrative virtual volume from said host computer;
13 reading said updated information from said
14 administrative virtual volume to update status information at
15 said virtual tape drive server.

1/6

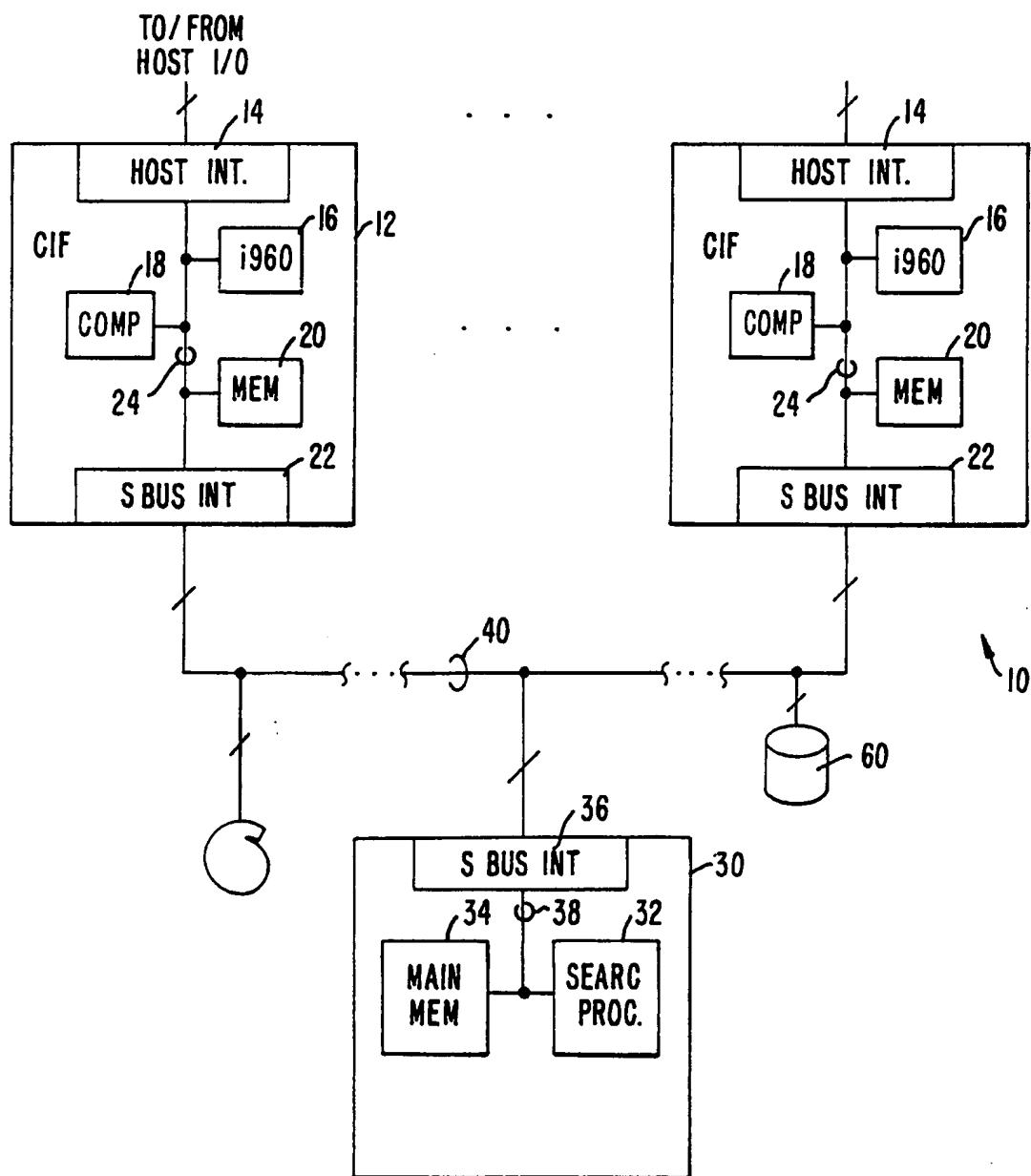


FIG. 1A.

2/6

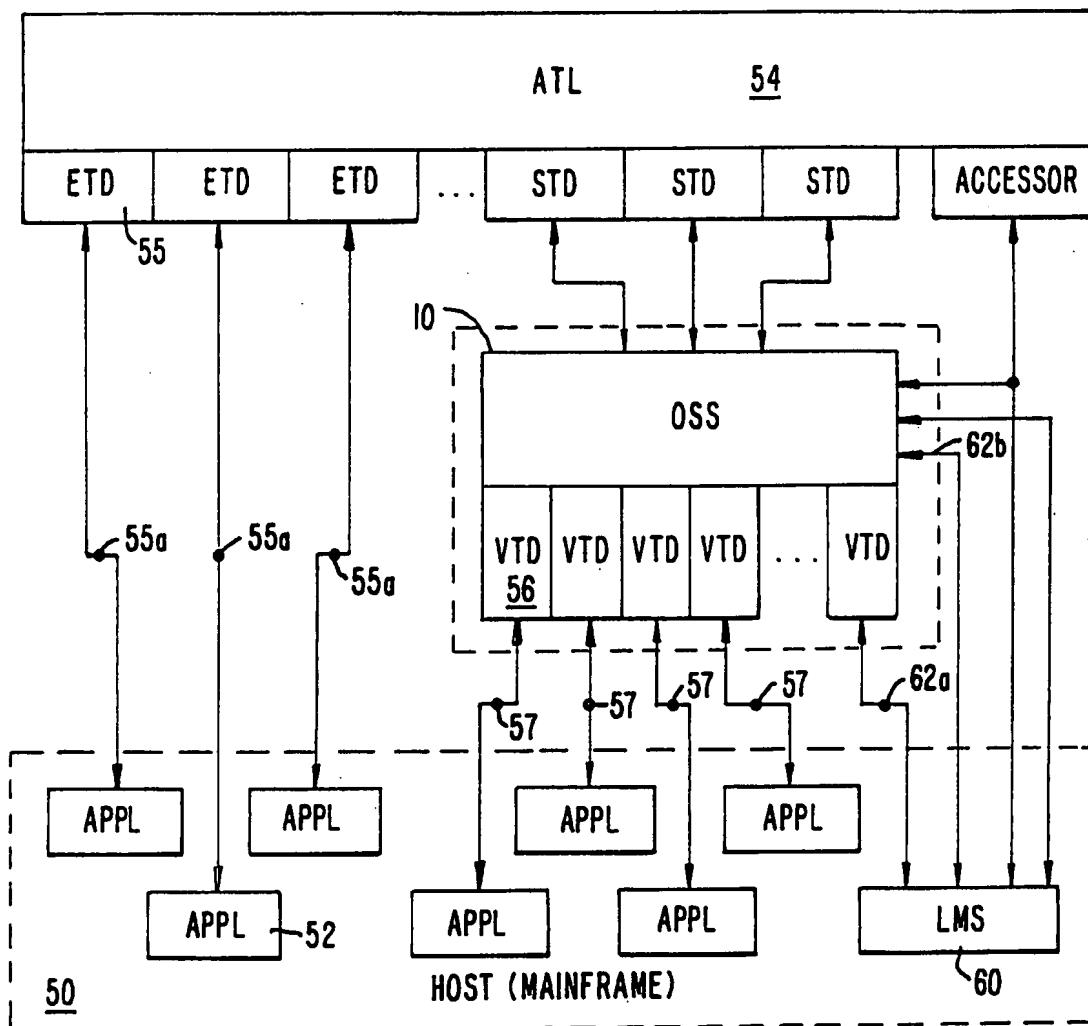


FIG. 1B.

3/6

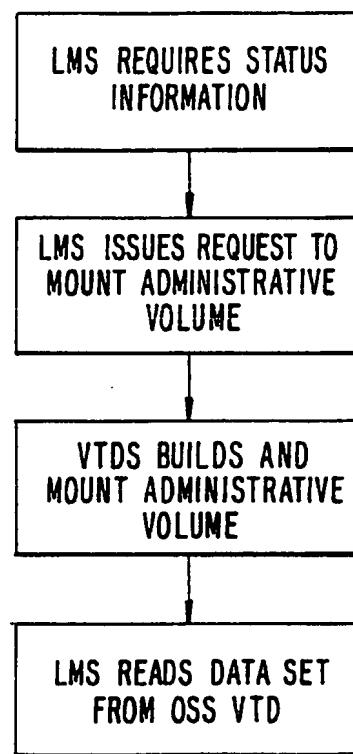
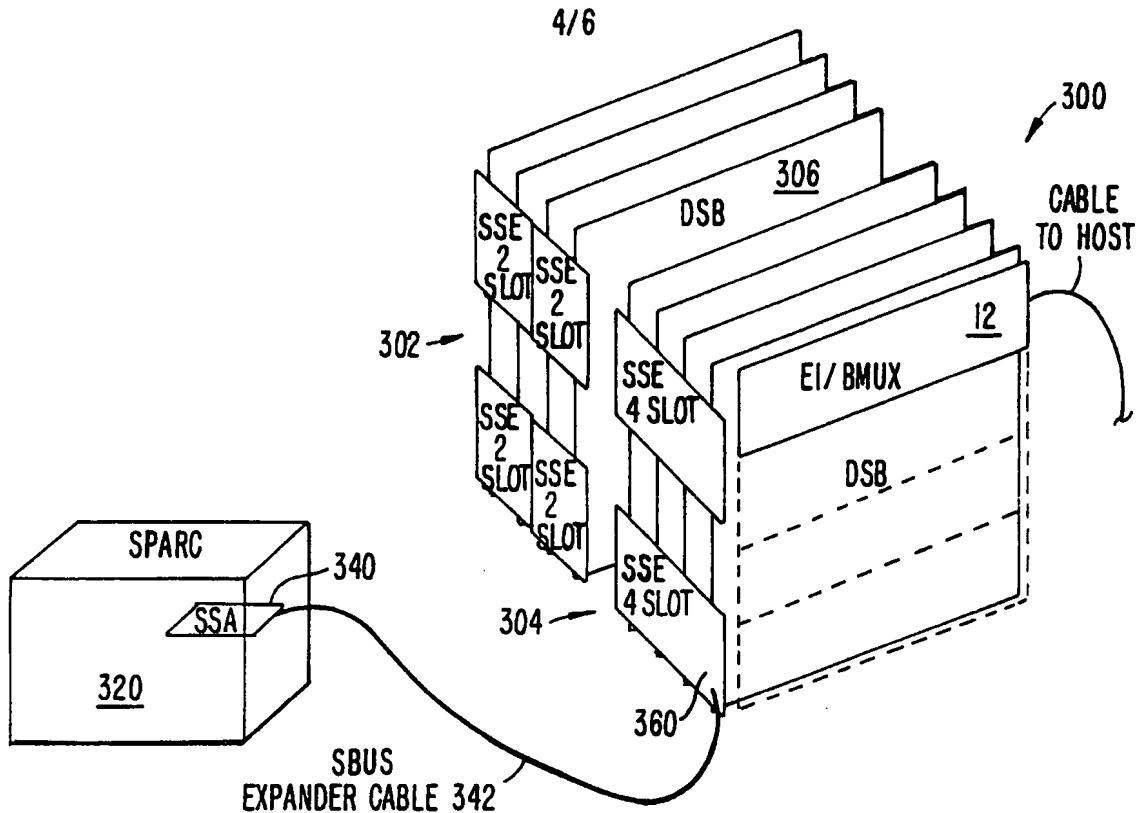
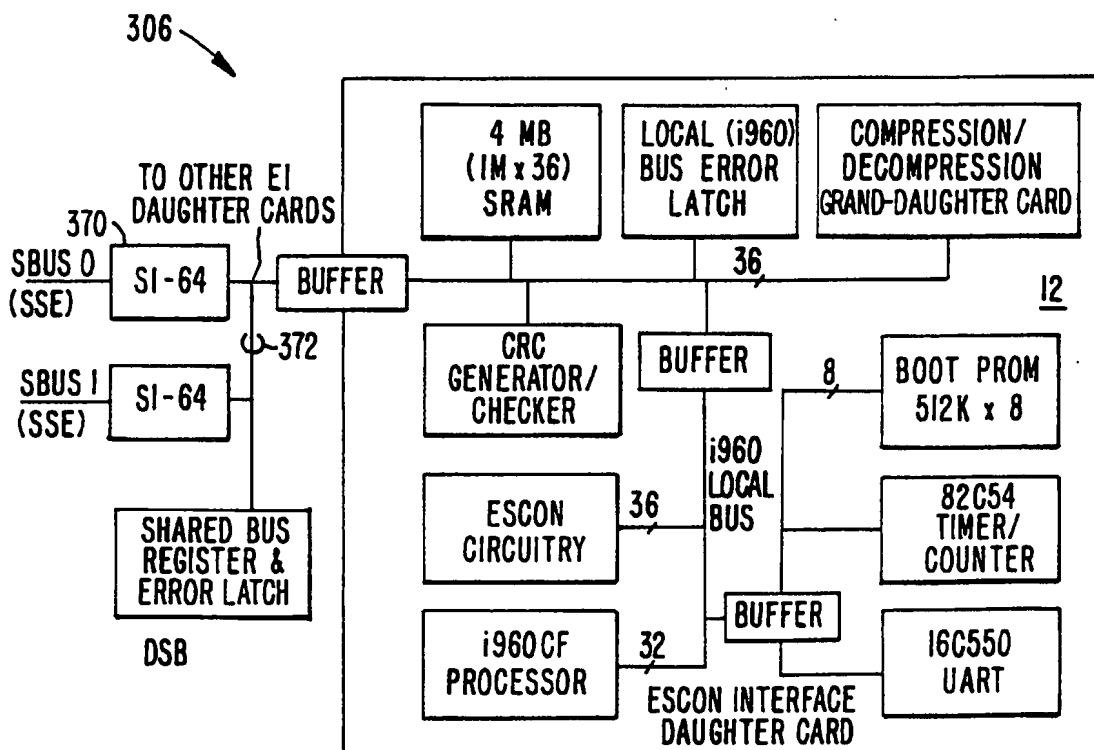


FIG. 2.

*FIG. 3.**FIG. 4.*

5/6

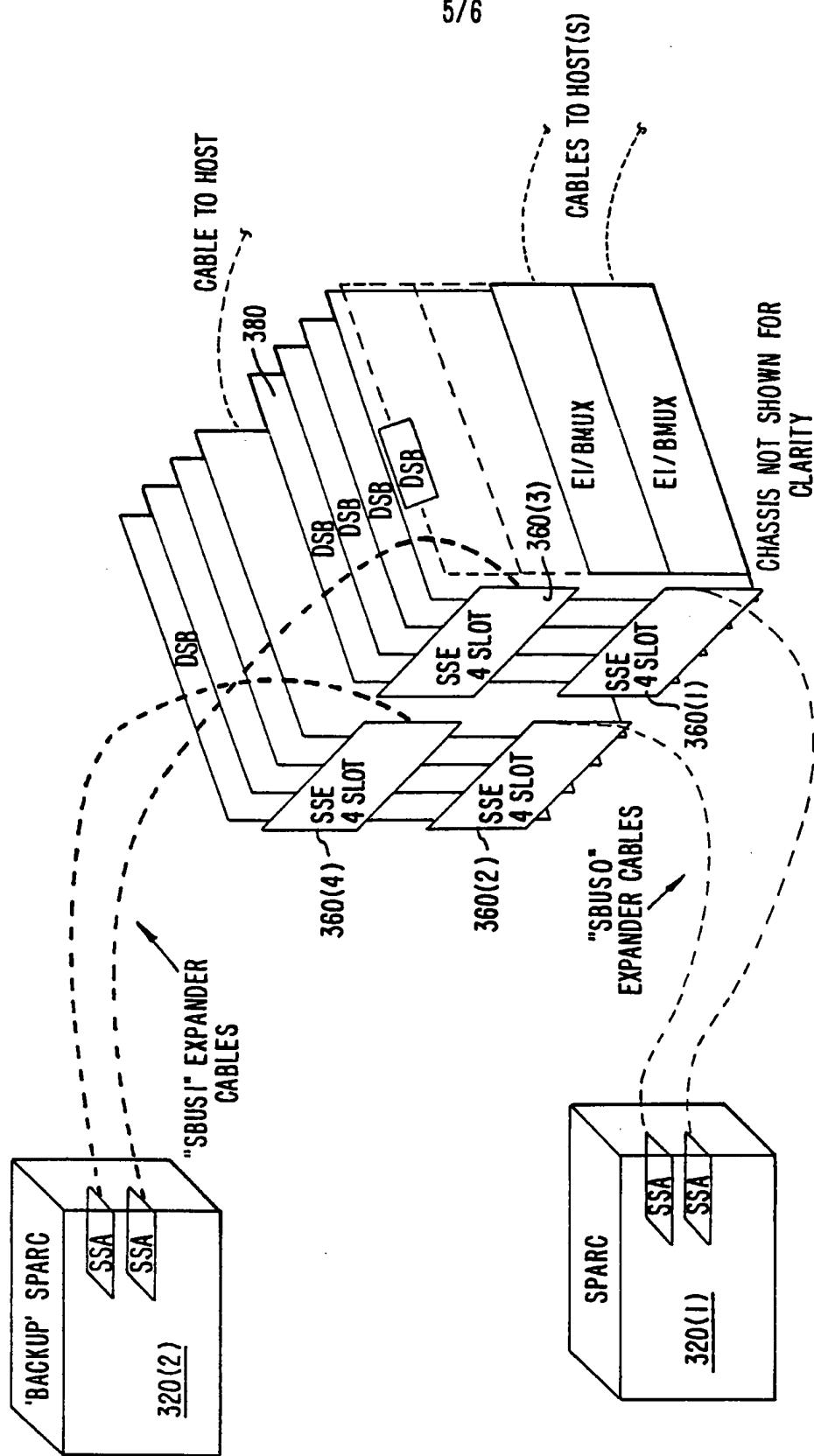


FIG. 5.

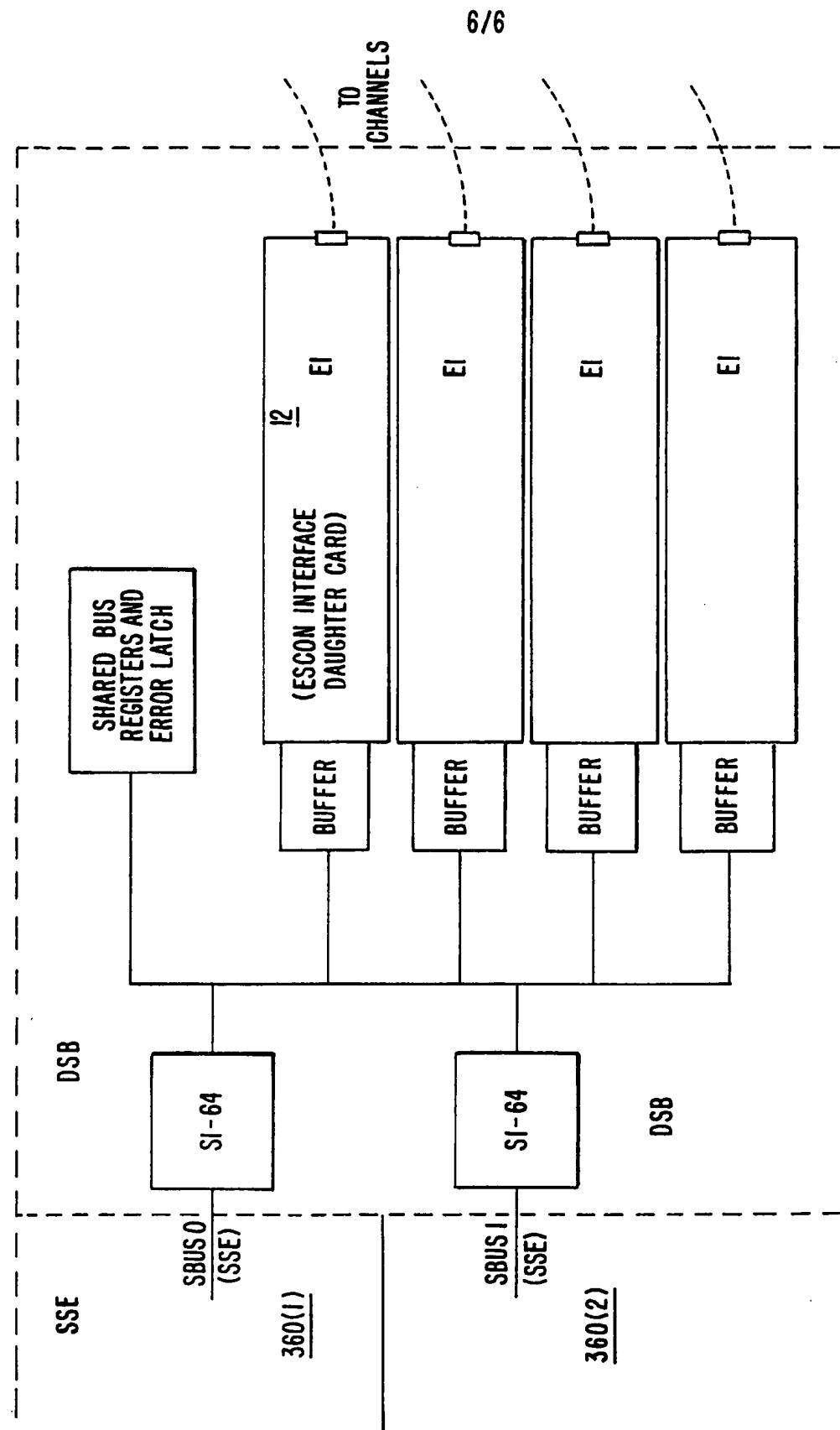


FIG. 6.